



ITS

Institut
Teknologi
Sepuluh Nopember



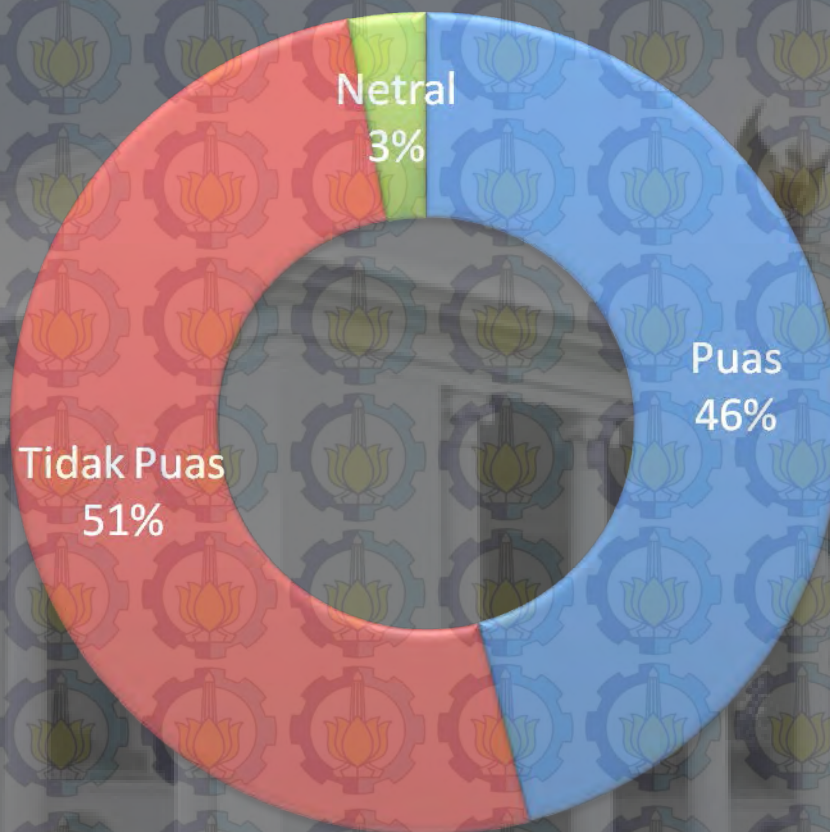
Analisis Microblogging untuk Menentukan Prioritas Kebijakan Publik Berdasarkan Opini Masyarakat dengan Menggunakan Algoritma Naïve Bayes dan Analytical Hierarchy Process

Dosen Pembimbing
Dr. Surya Sumpeno, S.T., M.Sc.
Dr. Adhi Dharma Wibawa, S.T., M.T.

Mohammad Khoiron
2214206708



Survey Tingkat Kepuasan Masyarakat Terhadap Kinerja Pemerintah (Oktober 2015)



Sumber:
Indobarometer
Oktober, 2015

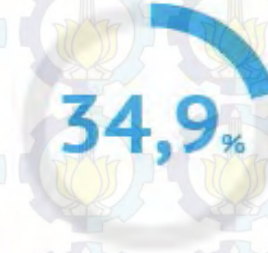
JUMLAH
PENDUDUK



JUMLAH PENGGUNA
INTERNET



PENETRASI



Jumlah Pengguna Internet
berdasarkan wilayah di Indonesia

SUMATERA



KALIMANTAN



SULAWESI



5.9
Juta

NUSA TENGGARA,
PAPUA DAN MALUKU

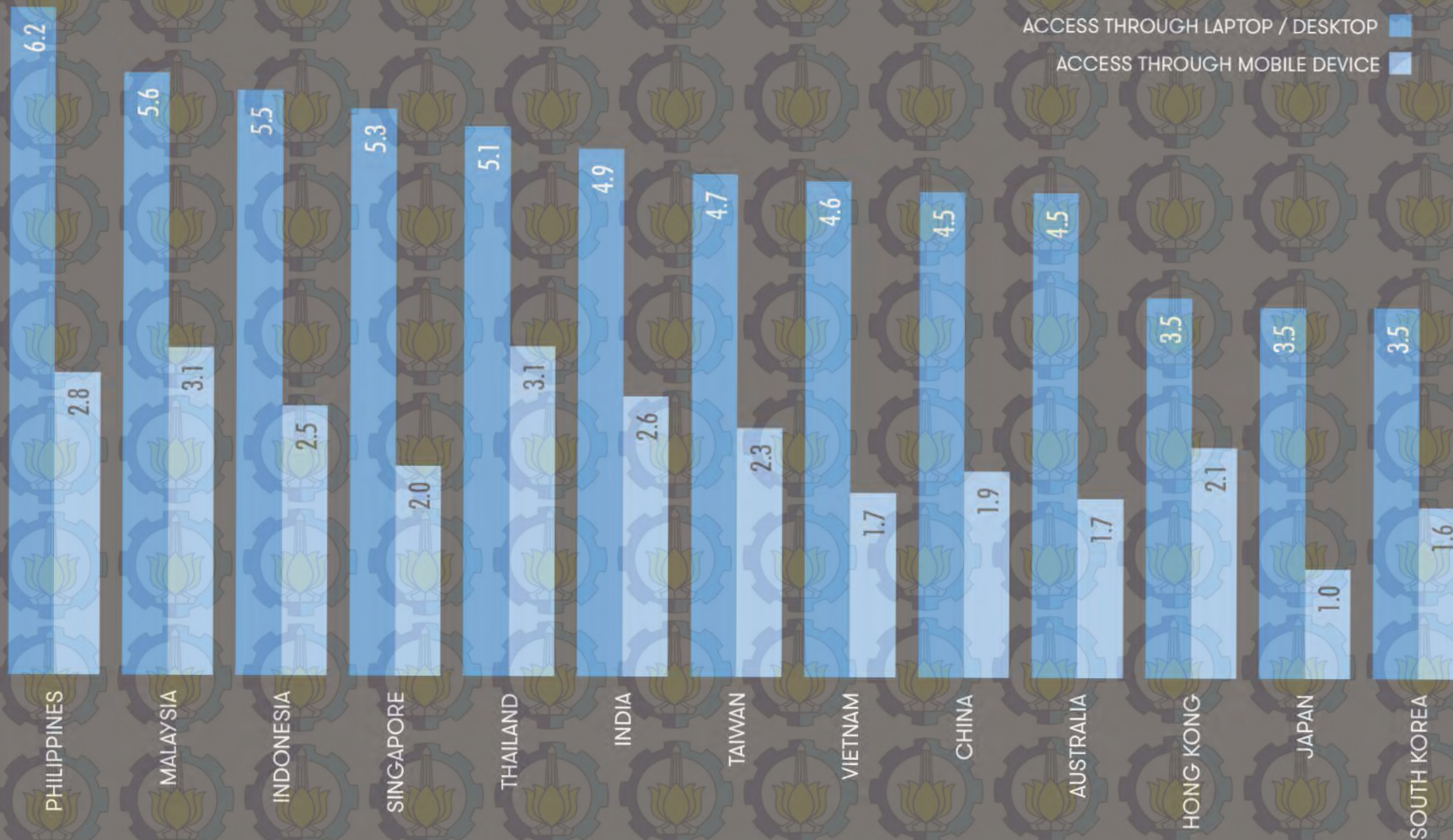
JAWA BALI



JAN
2014

TIME SPENT ON THE INTERNET

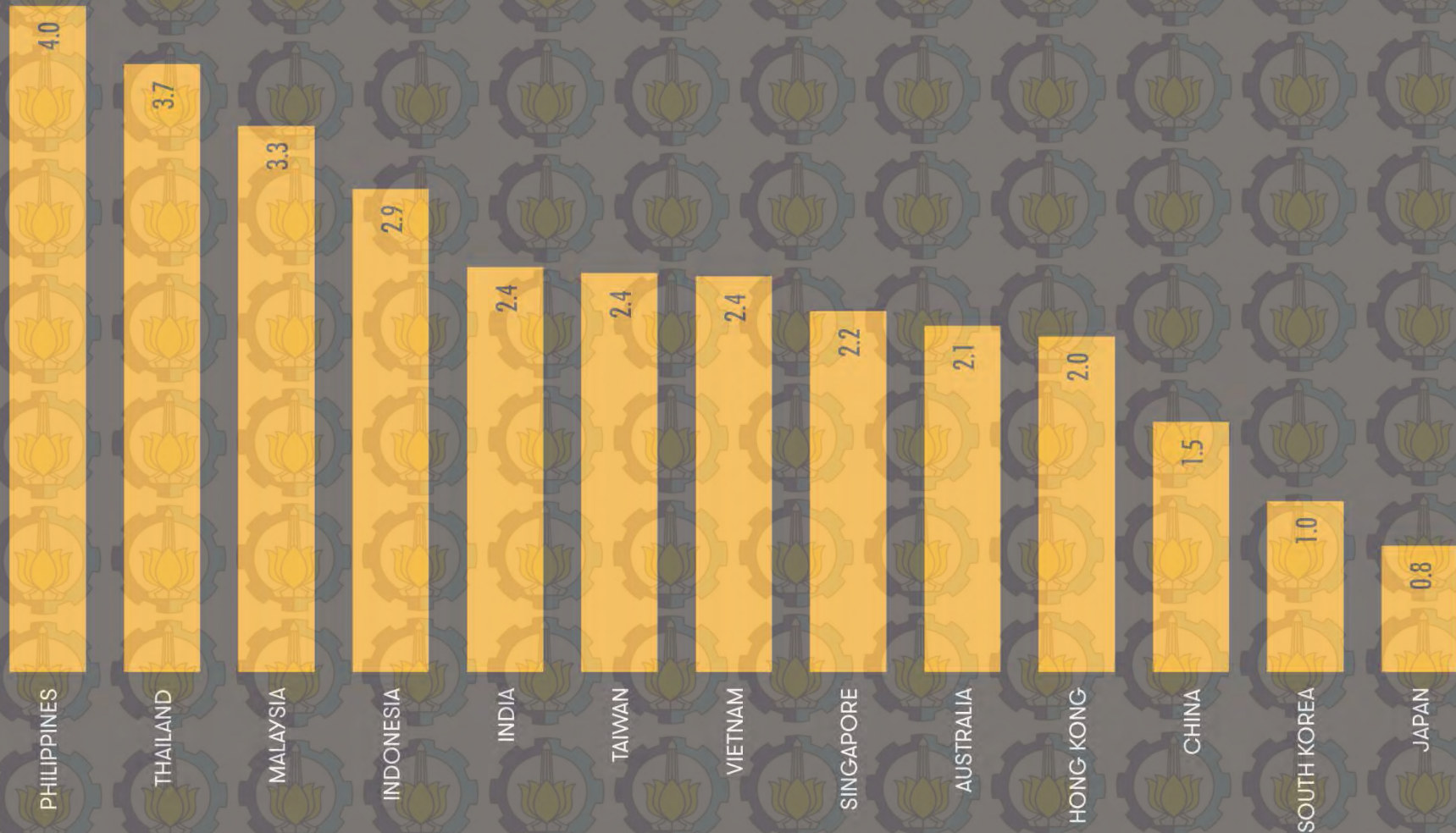
AVERAGE NUMBER OF HOURS PER DAY SPENT BY INTERNET USERS ON THE INTERNET



JAN
2014

TIME SPENT ON SOCIAL MEDIA

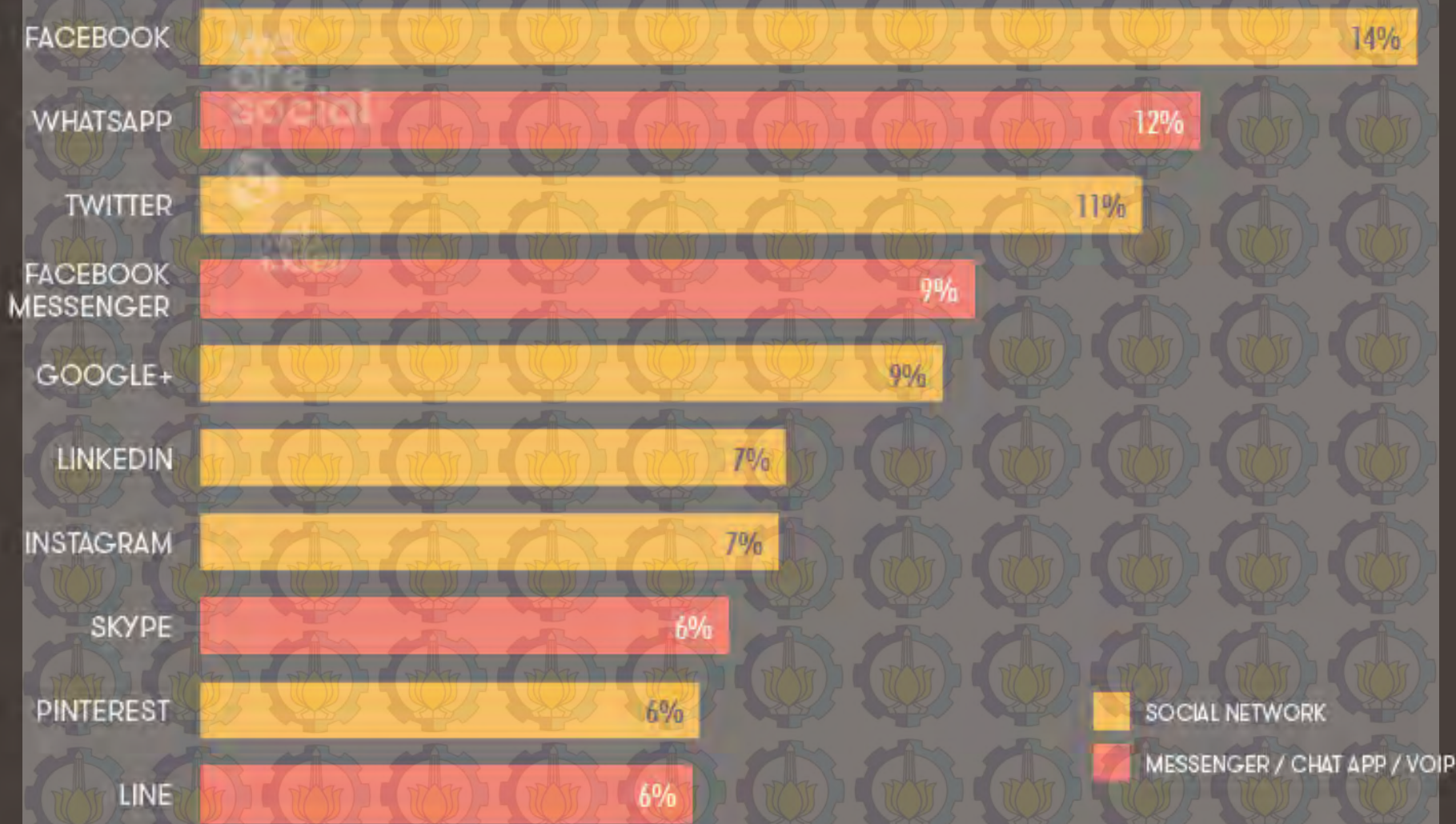
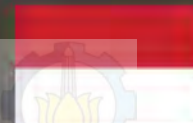
AVERAGE NUMBER OF HOURS PER DAY SPENT BY SOCIAL MEDIA USERS ON ALL SOCIAL CHANNELS



JAN
2015

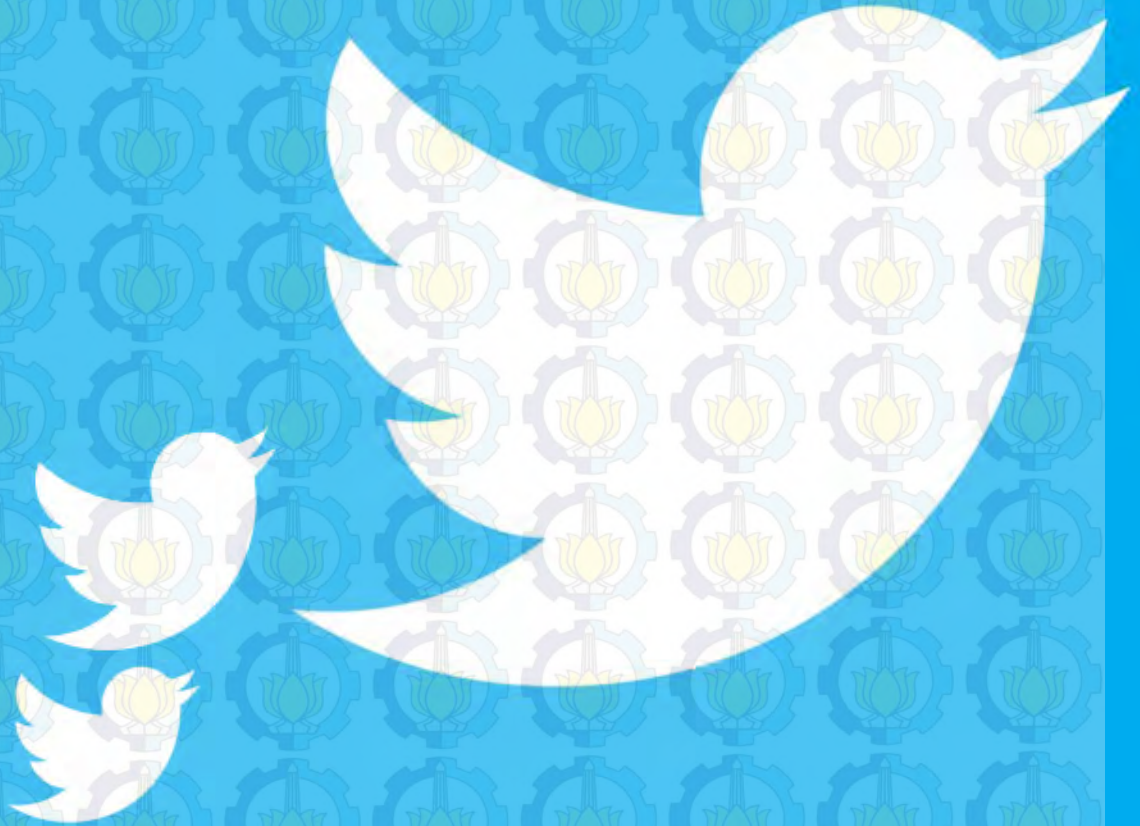
TOP ACTIVE SOCIAL PLATFORMS

SURVEY-BASED DATA: FIGURES REPRESENT USERS' OWN CLAIMED / REPORTED ACTIVITY



385
TWEETS

Dikirim
per Detik
Dari Indonesia



BAPPENAS

Menentukan kebijakan prioritas

Dengan melihat tuntutan masyarakat melalui survei dan analisis pendapat

Menentukan implementasi kebijakan

Dengan melihat database dan riset yang ada

Melihat progres suatu isu dan respon masyarakat

Melalui pembuatan sistem komplain dan analisis pendapat

#Permasalahan

Adanya kesulitan untuk memperoleh prioritas kebijakan publik yang mendekati aspirasi dan harapan masyarakat dengan memanfaatkan analisa terhadap microblogging

#Tujuan

Menghasilkan sebuah analisa dari microblogging, yang layak untuk menjadi bahan pertimbangan dan kajian untuk menentukan prioritas sebuah kebijakan publik berdasarkan aspirasi dan keinginan masyarakat.

#Manfaat

Meningkatnya kualitas kebijakan publik yang diambil oleh Pemerintah, sehingga dapat memenuhi ekspektasi dan harapan masyarakat yang berujung pada pelayanan prima

Penentuan kebijakan publik yang akan dianalisa



Penentuan Keyword yg mewakili kebijakan publik



Pengumpulan data (harvesting) dan penyimpanan di database



Pembelajaran dan klasifikasi (naïve bayes)



Feature selection (stopword removal dan stemming)



Pra proses (cleansing, case folding dan parsing)



Validasi dan evaluasi



Penentuan prioritas kebijakan (AHP)



Menyajikan hasil analisis untuk prioritas kebijakan publik

METHOD
OLOGY



Kata Kunci untuk Millenium Development Goals (MDGs)

No	Millenium Development Goals	Kata Kunci
1.	Memberantas Kemiskinan dan Kelaparan Ekstrem	Kemiskinan Kelaparan
2.	Mewujudkan Pendidikan Dasar untuk Semua	Pendidikan Buta huruf
3.	Mendorong Kesetaraan Gender dan Pemberdayaan Perempuan	Kesetaraan gender pemberdayaan perempuan
4.	Menurunkan Angka Kematian Anak	Kematian bayi Imunisasi
5.	Meningkatkan Kesehatan Ibu	Kesehatan ibu Kesehatan reproduksi
6.	Memerangi HIV dan AIDS Malaria Serta Penyakit Lainnya	Cegah hiv Cegah penyakit
7.	Memastikan Kelestarian Lingkungan	Keanekaragaman hayati Kelestarian lingkungan
8.	Mengembangkan Kemitraan Global untuk Pembangunan	Akses internet Perdaganganbebas

Tweet Harvesting

Mengumpulkan data tweet dengan parameter dan kata kunci yang telah ditentukan dengan menggunakan Tags V.6.0 & Twitter Streaming APIs

imunisasi

File Edit View Insert Format Data Tools Add-ons Help TAGS Last edit was made seconds ago by Mohamima

€ % 0.00 123 - Verdana 18 B I S A

TAGS v6.0

NS - New Sheets

Created by mhawksey. Read more about this at: <http://tags.hawksey.info>

With this spreadsheet you can:

- automatically pull results from a Twitter Search into a Google Spreadsheet

Instructions:

1. If there is no TAGS menu click this button -> [Enable custom menu](#)
2. If you've never run TAGS -> Setup Twitter Access do so now (this should only need be done once for all your TAGS sheets)
3. Enter term -> you can use serach operators like AND OR as well as from: and to: eg #JobsNow AND from:BarackObama' (without quotes)
4. Make a one off collection with TAGS -> Run now! or set a trigger to collect every hour TAGS -> Update archive every hour. To change the frequency open Tools -> Script Editor then Triggers -> Current scripts triggers... and adjust

Advanced Settings:

[+ Readme/Settings](#) [Archive](#)

Sentiment Analysis
Tweet Harvesting

Data Training

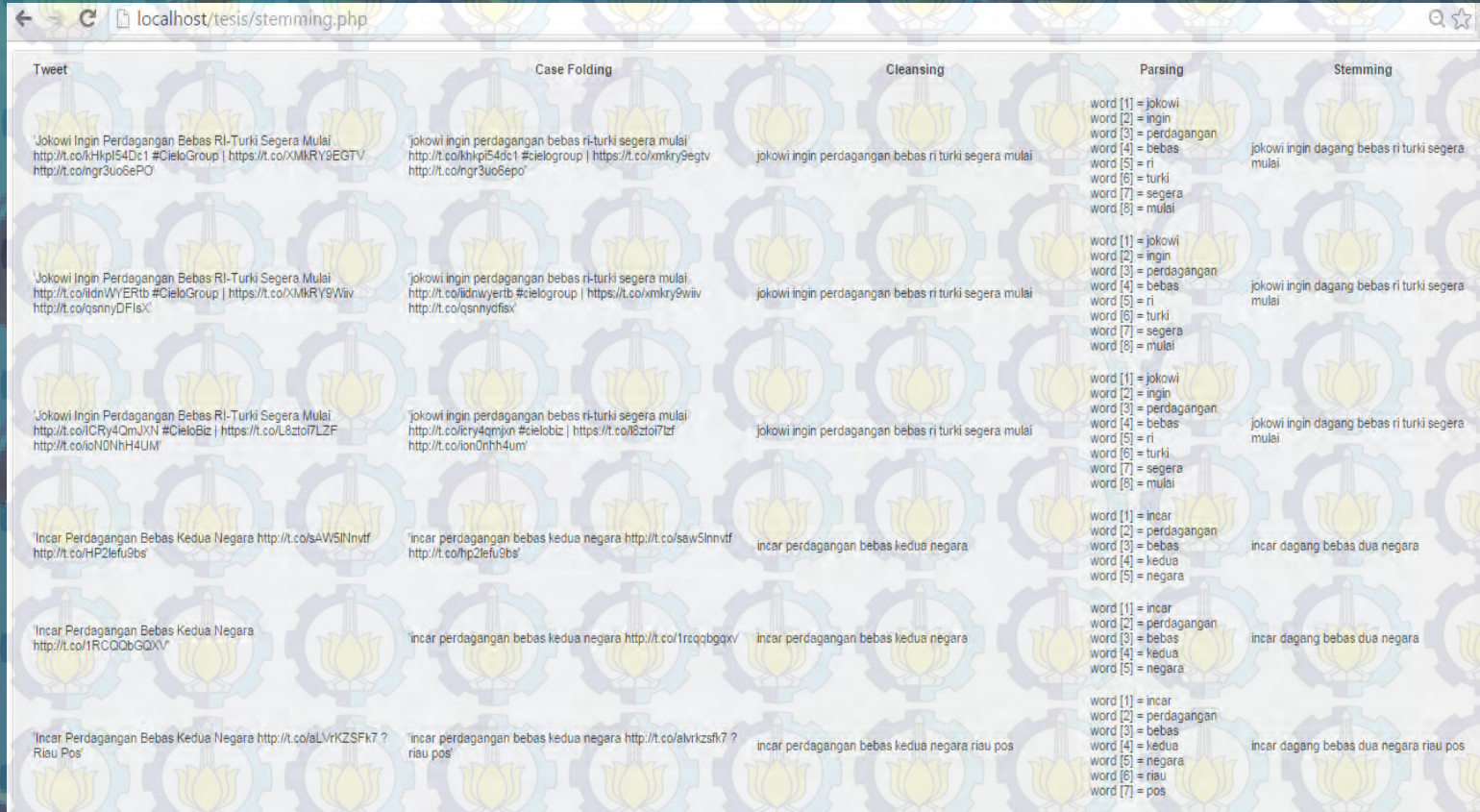
Dari hasil tweet harvesting yang telah dilakukan pada proses pengumpulan data, diambil 3200 tweet yang akan digunakan sebagai data training dan testing

J.t	Tweet ID	User ID	Tweet	Create	Kata Kunci	Klasifikasi
1	626454157239582720	gemuruh	Ujian Baru GunungMeletus & Kemarau. Itu blm sbpr, klu sdh kelaparan, Kematian. pemerintah msh woless??? :)	29/07/2015 19:08:04	Kelaparan	Negatif
2	624601340132462593	Marina_r	Semoga pemerintah tak membiarkan rakyatnya kelaparan, kasihan petani butuh uluran tangan pemerintah @Pak_JK @jokowi https://t.co/Oki	24/07/2015 16:25:38	Kelaparan	Negatif
3	624196759833567232	PartaiPe 4.	Pemerintah harus bisa mewujudkan zero hunger. Jangan ada lagi yang kelaparan di negeri yang berkelimpahan ini Ketum Bpk @Hary_	23/07/2015 13:37:59	Kelaparan	Negatif
4	620726882359644161	MUHAMMID @tempodotco	Soal kelaparan warga disana sudah biasa. Pemerintah saja yg nggak tanggap.	13/07/2015 23:49:55	Kelaparan	Negatif
5	620420923485634064	charlesh	Tolak HAMAS, Langkah Pemerintah Dipuji PDIP http://t.co/ND0rDqVLoB Tiga kampung di Papua terancam kelaparan http://t.co/XlIKVrjX3t	13/07/2015 03:34:09	Kelaparan	Negatif
6	620067856147877889	pritada	Dia menduga pemerintah daerah setempat menutup-nutupi warganya yang mengalami kelaparan. DPRD said.Citra lebih penting nyampe tir	12/07/2015 04:11:11	Kelaparan	Negatif
7	620059194310291456	pritada	NTT kelaparan sejak awal Januari 2015, pemerintah apatis, kemana aja????? Hei presiden @jokowi	12/07/2015 03:36:46	Kelaparan	Negatif
8	620039470629269504	pujiyuki	kondisi perekonomian Indonesia di bawah pemerintahan Presiden Joko Widodo memburuk.banyak kelaparan lapangan kerja modal tidak	12/07/2015 02:18:24	Kelaparan	Negatif
9	619034500236902400	risalno81	Masyarakat Timor Tengah Selatan, NTT kelaparan mn pemerintah? @jokowi @setkabgoid	09/07/2015 07:45:00	Kelaparan	Negatif
10	618355310080622592	widjidiot	Makan? Itu barang mewah!! Kelaparan dimana2. Kaum pejuang semena2, pemerintahan tidak ada gunanya.	07/07/2015 10:46:09	Kelaparan	Negatif
11	618352093460443137	TitilMelia	Resafull kabinet ga perlu, presiden nya yg hrs di resafull... kl pemerintah ini ampe 5 thn bisa mati kelaparan kt semua	07/07/2015 10:33:22	Kelaparan	Negatif
12	618319201481048064	ElyGoroL	Pemerintah daerah Kab/Kota d NTT memang kurang gizi makanya rakyatnya gizi buruk dan kelaparan!	07/07/2015 08:22:40	Kelaparan	Negatif
13	617620460239323136	omellin	RT @nayasaa: Cth pemerintahan ga ada kerjaan, raja restoran yg buka selama puasa. Tpi ga pernah ngok warganya kelaparan. Eshol.	05/07/2015 10:06:07	Kelaparan	Negatif
14	617379513027067908	DNuhdia	Pemerintah: tenang kita akan kuat semesta pangan tahun 2018... rakyat : kalau begitu kami akan kelaparan duluan.	04/07/2015 18:08:40	Kelaparan	Negatif
15	617222396198531072	ghosyis	RT @IndoPolitik1: busung lapar dan orang mati kelaparan karena tidak bisa beli makanan akan meningkat, apabila pemerintah tidak melak	04/07/2015 07:44:21	Kelaparan	Negatif
16	617014270929317888	arekmal	WAHA! PEMERINTAH DAN PEJABAT JANGAN MENGURUSI DIRI ANDA, TAPI LUPA KEPADA RAKYAT YG KELAPARAN	03/07/2015 17:57:20	Kelaparan	Negatif
17	616856449138888705	bahagia_	@kompascom pemerintah sekarang memang gak ada solusi dr kelaparan ntt, sinabung, pssi dan rupiah loyo, semuanya tida ada solusi	03/07/2015 07:30:12	Kelaparan	Negatif
18	616852776123957248	ely_susi_	@AnakBaranusa @NdoroTweet Sy cc kan tweet ndoro ke anda. Itu krn berhari2 anda teriak2 ttg kelaparan di NTT dn pemerintah tdk bereaks	03/07/2015 07:15:37	Kelaparan	Negatif
19	616034109777817601	Movin_br	Begini Perjuangan Warga Kelaparan Bertahan Hidup: Mereka menagih janji pemerintah soal penyediaan sumur b... http://t.co/C0sWj3WHbD	01/07/2015 01:02:31	Kelaparan	Negatif
20	615786178202071040	lifeology	Rakyat yg kelaparan terombang ambing dalam peperangan korporasi dgn pemerintah	30/06/2015 08:37:20	Kelaparan	Negatif
21	615780913427812352	bang_sy 11.	D benak pemerintah, kalau subsidi dicabut, akan ada uang buat pembangunan. Bener sih, tp rakyat mati kelaparan duluan. #SudahlahJ	30/06/2015 08:16:25	Kelaparan	Negatif
22	615378943571443712	akhyar98	Separah itukah jokower, menghubungkan kasus kelaparan NTT yang menewaskan 11 anak ke DPR dan bukannya pemerintah?	29/06/2015 05:39:07	Kelaparan	Negatif

Sentiment Analysis
Data Training

Pra Proses dan Seleksi Fitur

Mengumpulkan data tweet dengan parameter dan kata kunci yang telah ditentukan dengan menggunakan Twitter Streaming APIs



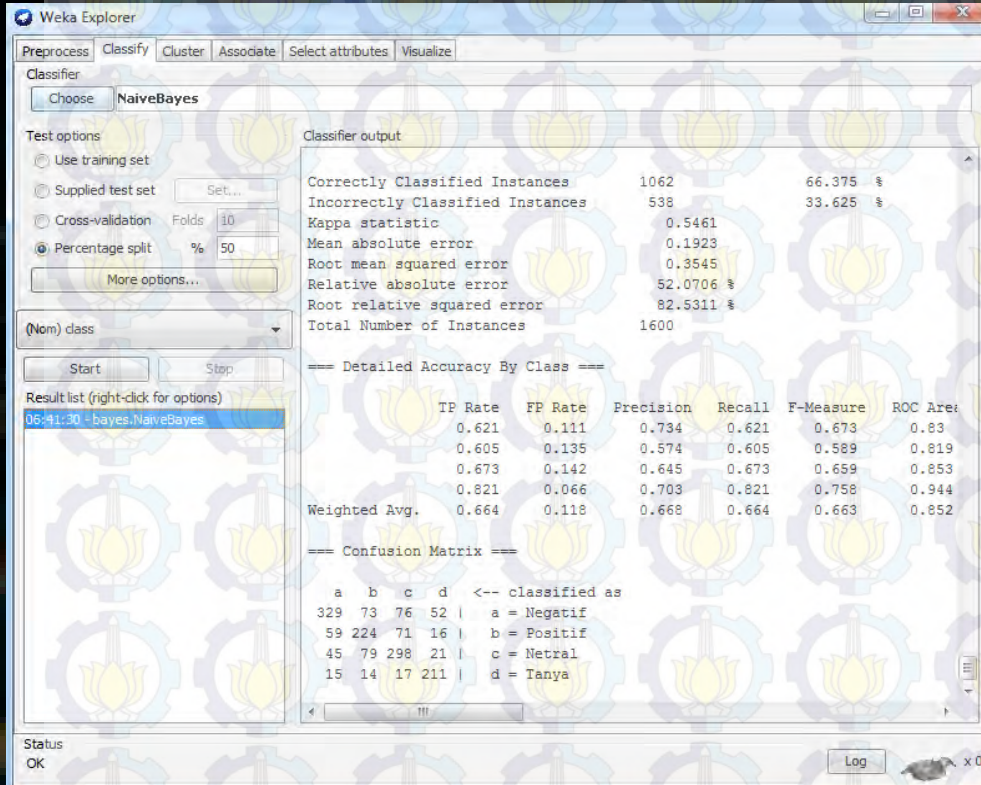
Tweet	Case Folding	Cleansing	Parsing	Stemming
'Jokowi Ingin Perdagangan Bebas RI-Turki Segera Mulai' #CieloGroup https://t.co/XMKRY9EGTV http://t.co/ngr3uo6ePO	'jokowi ingin perdagangan bebas ri-turki segera mulai' #cielogroup https://t.co/xmkry9egtv http://t.co/ngr3uo6epo	jokowi ingin perdagangan bebas ri turki segera mulai	word [1] = jokowi word [2] = ingin word [3] = perdagangan word [4] = bebas word [5] = ri word [6] = turki word [7] = segera word [8] = mulai	jokowi ingin dagang bebas ri turki segera mulai
'Jokowi Ingin Perdagangan Bebas RI-Turki Segera Mulai' #CieloGroup https://t.co/XMKRY9Wiv http://t.co/qsnnyDFIsX	'jokowi ingin perdagangan bebas ri-turki segera mulai' #cielogroup https://t.co/xmkry9wiv http://t.co/qsnnydfisx	jokowi ingin perdagangan bebas ri turki segera mulai	word [1] = jokowi word [2] = ingin word [3] = perdagangan word [4] = bebas word [5] = ri word [6] = turki word [7] = segera word [8] = mulai	jokowi ingin dagang bebas ri turki segera mulai
'Jokowi Ingin Perdagangan Bebas RI-Turki Segera Mulai' #CieloBiz https://t.co/L8zto7LZF http://t.co/lonDNhH4UM	'jokowi ingin perdagangan bebas ri-turki segera mulai' #cielobiz https://t.co/l8zto7lzf http://t.co/lonDNhH4um	jokowi ingin perdagangan bebas ri turki segera mulai	word [1] = jokowi word [2] = ingin word [3] = perdagangan word [4] = bebas word [5] = ri word [6] = turki word [7] = segera word [8] = mulai	jokowi ingin dagang bebas ri turki segera mulai
'Incar Perdagangan Bebas Kedua Negara http://t.co/sAW5lNnvf http://t.co/HP2lefU9bs	'incar perdagangan bebas kedua negara http://t.co/saw5lnnvf http://t.co/HP2lefU9bs	incar perdagangan bebas kedua negara	word [1] = incar word [2] = perdagangan word [3] = bebas word [4] = kedua word [5] = negara	incar dagang bebas dua negara
'Incar Perdagangan Bebas Kedua Negara' http://t.co/1RCQQbGQXV	'incar perdagangan bebas kedua negara http://t.co/1rcqqbgqxv	incar perdagangan bebas kedua negara	word [1] = incar word [2] = perdagangan word [3] = bebas word [4] = kedua word [5] = negara	incar dagang bebas dua negara
'Incar Perdagangan Bebas Kedua Negara http://t.co/aLvrKzSFk7 ? Riau Pos	'incar perdagangan bebas kedua negara http://t.co/aLvrKzSFk7 ? riau pos	incar perdagangan bebas kedua negara riau pos	word [1] = incar word [2] = perdagangan word [3] = bebas word [4] = kedua word [5] = negara word [6] = riau word [7] = pos	incar dagang bebas dua negara riau pos

Sentiment Analysis

Pre Process (cleansing, case folding, parsing) and feature Selection (stemming)

Pembelajaran dan Klasifikasi

Proses pembelajaran dan klasifikasi memanfaatkan software open source WEKA V. 3.6



The screenshot displays the Weka Explorer window with the NaiveBayes classifier selected. The 'Classifier output' tab shows the following results:

Metric	Value	Percentage
Correctly Classified Instances	1062	66.375 %
Incorrectly Classified Instances	538	33.625 %
Kappa statistic	0.5461	
Mean absolute error	0.1923	
Root mean squared error	0.3545	
Relative absolute error	52.0706 %	
Root relative squared error	82.5311 %	
Total Number of Instances	1600	

Below the main results, the 'Detailed Accuracy By Class' table is shown:

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
0.621	0.111	0.734	0.621	0.673	0.83	
0.605	0.135	0.574	0.605	0.589	0.819	
0.673	0.142	0.645	0.673	0.659	0.853	
0.821	0.066	0.703	0.821	0.758	0.944	
Weighted Avg.	0.664	0.118	0.668	0.664	0.663	0.852

The 'Confusion Matrix' is also displayed:

	a	b	c	d	<-- classified as
329	73	76	52		a = Negatif
59	224	71	16		b = Positif
45	79	298	21		c = Netral
15	14	17	211		d = Tanya

The status bar at the bottom indicates 'OK' and 'Log'.

Sentiment Analysis
Training and Clasification Process

Hasil Proses Pembelajaran

Algoritma Naïve Bayes

Untuk mendapatkan akurasi terbaik dilakukan variasi pada nilai term weighting, min term freq, stemmer, stop word, tokenizer dan percentage split.

Term Weighting	Min Term Freq	Stemmer	Stop Word	Tokenizer	Percentage Split	Accuracy
TF-IDF	3	Lovins	-	Word	66%	72,2%
TF-IDF	3	Nazief-Andriani	-	Word	66%	67,7%

Algoritma SVM-SMO

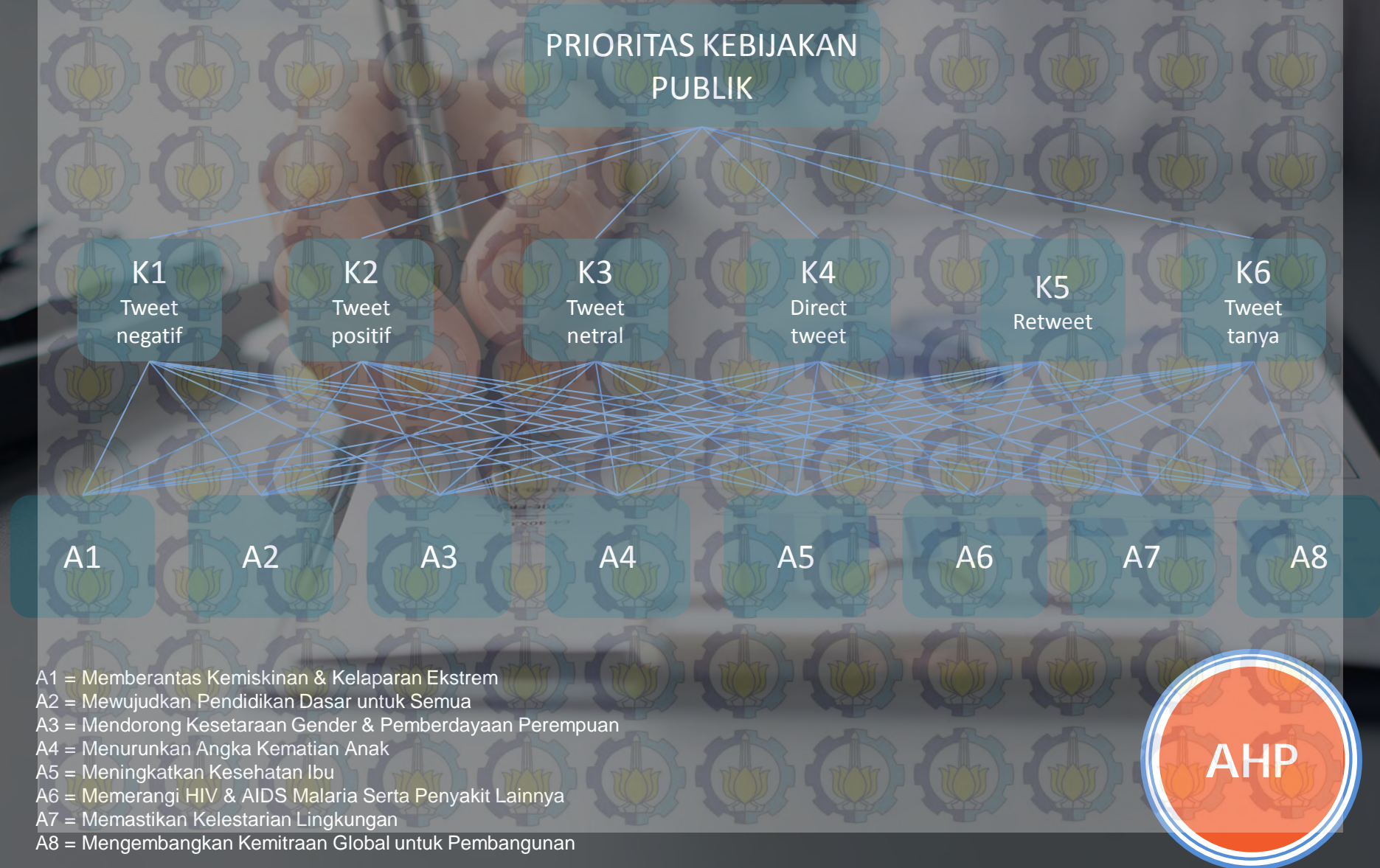
Untuk mendapatkan akurasi terbaik dilakukan variasi Build Logistic Model, C, Epsilon, Kernel, Num of folds, Random Seed, Tolerance Parameter, Procentage Split.

Build Logistic Model	C	Epsilon	Kernel	Num of folds	Random Seed	Tolerance Parameter	Procentage Split (%)	Accuracy
TRUE	3	0,00001	RBF	-1	1	0,001	80	70,63%

Sentiment Analysis

Result of Training and Clasification Process

Pembentukan Hirarki



Perbandingan Berpasangan

KRITERIA

- a) Jumlah tweet negatif sedikit lebih penting dari jumlah tweet positif.
- b) Jumlah tweet negatif sedikit lebih penting dari jumlah tweet netral.
- c) Jumlah tweet negatif sedikit lebih penting dari jumlah direct tweet.
- d) Jumlah tweet negatif sedikit lebih penting dari jumlah retweet.
- e) Jumlah tweet negatif sedikit lebih penting dari jumlah tweet tanya.
- f) Jumlah tweet positif sedikit lebih penting dari jumlah tweet netral
- g) Jumlah direct tweet sedikit lebih penting dari jumlah tweet tanya

Kriteria	Jumlah Tweet Negatif	Jumlah Tweet Positif	Jumlah Tweet Netral	Jumlah Direct Tweet	Jumlah Retweet	Jumlah Tweet Tanya
Jumlah Tweet Negatif	1,0000	3,0000	3,0000	3,0000	3,0000	3,0000
Jumlah Tweet Positif	0,3333	1,0000	3,0000	1,0000	1,0000	1,0000
Jumlah Tweet Netral	0,3333	0,3333	1,0000	1,0000	1,0000	1,0000
Jumlah Direct Tweet	0,3333	1,0000	1,0000	1,0000	3,0000	1,0000
Jumlah Retweet	0,3333	1,0000	1,0000	0,3333	1,0000	1,0000
Jumlah Tweet Tanya	0,3333	1,0000	1,0000	1,0000	1,0000	1,0000
Jumlah	2,6667	7,3333	10,0000	7,3333	10,0000	8,0000

AHP

Matrik Perbandingan Berpasangan (Saaty, L Thomas)

Normalisasi Matrik Perbandingan Berpasangan

Kriteria	Jumlah Tweet Negatif	Jumlah Tweet Positif	Jumlah Tweet Netral	Jumlah Direct Tweet	Jumlah Retweet	Jumlah Tweet Tanya	Jumlah	Bobot
Jumlah Tweet Negatif	0,3750	0,4091	0,3000	0,4091	0,3000	0,3750	2,1682	0,3614
Jumlah Tweet Positif	0,1250	0,1364	0,3000	0,1364	0,1000	0,1250	0,9227	0,1538
Jumlah Tweet Netral	0,1250	0,0455	0,1000	0,1364	0,1000	0,1250	0,6318	0,1053
Jumlah Direct Tweet	0,1250	0,1364	0,1000	0,1364	0,3000	0,1250	0,9227	0,1538
Jumlah Retweet	0,1250	0,1364	0,1000	0,0455	0,1000	0,1250	0,6318	0,1053
Jumlah Tweet Tanya	0,1250	0,1364	0,1000	0,1364	0,1000	0,1250	0,7227	0,1205

KETERANGAN:

membagi setiap nilai pada kolom matrik dengan hasil penjumlahan kolom yang bersesuaian pada matrik perbandingan berpasangan

AHP

Pemeriksaan Rasio Konsistensi (CR)

Pemeriksaan konsistensi dilakukan untuk melihat apakah matriks berpasangan yang sudah kita buat nilainya konsisten. Hal ini terpenuhi jika nilai $CR \leq 0.1$

Nilai Eigen Maksimum

$\lambda_{maks} = 6,2889$ (jumlah perkalian bobot pada normalisasi matrik dengan jumlah pada matrik berpasangan)

Nilai Konsistensi Index (CI)

$$CI = (\lambda_{maks} - n) / (n - 1)$$

$$CI = 0,057777778$$

Nilai Random Index (RI)

Nilai RI diambil dari tabel Random Index. Untuk Matriks yang mempunyai ordo 6 nilainya adalah $= 1,24$

Nilai Consistency Ratio (CR)

$$CR = CI / RI$$

$$CR = 0,046594982 \text{ (Nilai } CR \leq 0,1 \text{ sehingga konsisten)}$$

AHP

Evaluasi Bobot Keseluruhan

Kriteria/ Alternatif	Jumlah Tweet Negatif	Jumlah Tweet Positif	Jumlah Tweet Netral	Jumlah Direct Tweet	Jumlah Retweet	Jumlah Tweet Tanya	Nilai Akhir	Peringkat /Prioritas
	0,3614	0,1538	0,1053	0,1538	0,1053	0,1205		
A1	11950	22297	88383	84796	32048	8833	34533,63	1
A2	7198	15052	84160	73059	29444	5990	28835,87	2
A3	84	908	8088	8225	740	145	2381,98	5
A4	481	1304	11601	6890	3355	4659	3570,06	4
A5	154	1390	4777	4864	1115	368	1682,21	6
A6	76	1397	2954	2926	612	908	1177,17	7

- A1 = Memberantas Kemiskinan & Kelaparan Ekstrem
 A2 = Mewujudkan Pendidikan Dasar untuk Semua
 A3 = Mendorong Kesetaraan Gender & Pemberdayaan Perempuan
 A4 = Menurunkan Angka Kematian Anak
 A5 = Meningkatkan Kesehatan Ibu
 A6 = Memerangi HIV & AIDS Malaria Serta Penyakit Lainnya
 A7 = Memastikan Kelestarian Lingkungan
 A8 = Mengembangkan Kemitraan Global untuk Pembangunan

AHP

Prioritas Kebijakan Publik

Dari evaluasi bobot keseluruhan terlihat bahwa urutan atau prioritas kebijakan publik yang dapat diambil berdasarkan algoritma Analytical Hierarchy Process (AHP) adalah sebagai berikut:

1. Memberantas Kemiskinan dan Kelaparan Ekstrem
2. Mewujudkan Pendidikan Dasar untuk Semua
3. Mengembangkan Kemitraan Global untuk Pembangunan
4. Menurunkan Angka Kematian Anak
5. Mendorong Kesetaraan Gender dan Pemberdayaan Perempuan
6. Meningkatkan Kesehatan Ibu
7. Memerangi HIV dan AIDS Malaria Serta Penyakit Lainnya
8. Memastikan Kelestarian Lingkungan

AHP

Kesimpulan

1. Penelitian yang telah dilakukan berhasil membuktikan bahwa analisis microblogging dapat menjadi bahan pertimbangan dan kajian untuk menentukan prioritas kebijakan publik yang mendekati aspirasi dan keinginan masyarakat.
2. Akurasi tertinggi yang bisa dicapai dalam penelitian ini adalah sebesar 72,2 % dengan menggunakan algoritma Naïve Bayes, pembobotan yang digunakan adalah TF-IDF dengan minimal term frequency=3, tanpa melakukan proses stopwords, tokenizer yang dipakai adalah= word tokenizer dengan percentage split sebesar 66%.
3. Penggunaan stop word removal cenderung menurunkan nilai akurasi.

TERIMA KASIH

Mohammad Khoiron

Telematika-CIO
Institut Teknologi Sepuluh Nopember Surabaya
2015

Gracias Dziękuję Děkují danke Kiitos